APPENDIX 14

VIRGINIA MUNICIPAL STORMWATER ASSOCIATION

**COMMENTS OF THE
VIRGINIA MUNICIPAL STORMWATER ASSOCIATION, INC.
REGARDING U.S. EPA'S DRAFT CHESAPEAKE BAY TMDL AND
VIRGINIA'S DRAFT CHESAPEAKE BAY TMDL WIP**

MODELING COMMENTS

EPA expects VAMSA members (and others) to comply with an extraordinarily expensive and operationally cumbersome clean-up plan. However, EPA itself has not fulfilled its obligation to ensure that its modeling framework is adequate to support its TMDL and the accompanying WLAs and LAs. If EPA presses forward with finalizing the TMDL over the objections of Bay dischargers and interested stakeholders, despite the faulty model that it has put forth in support of its TMDL, its decision to do so will be arbitrary and capricious.

Like any model, EPA's Bay model is a highly imperfect representation of reality. Over time, EPA has inappropriately shifted to using it in ways that are beyond its capabilities (e.g., predicting D.O. concentrations and non-attainment rates in specific segments to the single percentage point level under far-reaching management scenarios). This has resulted in wide swings in predicted loads and goals with each major model version. VAMSA believes that this instability will continue to occur in the future as the model is periodically modified.

VAMSA objects to overreliance on unstable models to the single percentage point of output, such that environmental policies are undermined with each new model run. Following are examples of problematic modeling issues that should cause EPA to shy away from major disruptions to state regulations/policy on the basis of single-digit shifts in model output:

1. *Lack of full model validation and peer review*: The Scientific and Technical Advisory Committee (STAC) has placed a strong emphasis on the need for model validation (STAC, 2006), calling validation "an essential and a required step in model development, particularly if the model is to be used for TMDL development purposes" (STAC, 2008b).[1] Although the watershed model (WSM) appears to have been subjected to some kind of validation, the public documentation of the validation is very poor. Moreover, it is unclear if the Water

[1] Attachment A to this Appendix. Additional References Include: Scientific and Technical Advisory Committee. 2006. Modeling in the Chesapeake Bay Program: 2010 and Beyond. STAC Publication 06-001. 42 p.
Scientific and Technical Advisory Committee. 2008a. Review of the Phase V Community Chesapeake Bay Watershed Model. STAC Publication 08-003
Scientific and Technical Advisory Committee. 2008b. Second Review of the Phase V Community Chesapeake Bay Watershed Model. STAC Publication 08-003.

Quality and Sediment Transport Model (WQSTM) has been validated in any manner. It also appears that the STAC reviews of the WQSTM have focused on the sediment, clarity, and SAV components, and there may not have been a complete peer review of the latest version of the full eutrophication and D.O. simulation.

2. *The model is being extrapolated beyond the observed range of management controls and living resources*: The model framework has been calibrated using data from years with widely varying hydrologic conditions. However, none of the calibration data are representative of management controls or living resources that being called for as part of the Bay TMDL and related goals. Therefore, there is simply no way to verify that the Bay system will respond precisely as predicted. The model predictions of attainment are best characterized as rough approximations rather than highly precise predictions.

3. *An estimate of model uncertainty should be used to determine the essential equivalence of model scenarios*: EPA was correct to implement an interpretive rule (the "1% rule") by which model-predicted non-attainment is considered indistinguishable from zero. However, the one-percent magnitude underestimates the model error and overestimates the precision of both the model and monitoring data. Based on the analysis of Bell (2010b), segments that are close to attainment would require spatial D.O. violation rates that differ by 4% or more before they would be statistically distinguishable from one another. EPA's justification for the 1% magnitude was not based on calibration or validation statistics, but by an analysis of the sensitivity of simulated to DO attainment to simulated load reductions.[2] It is recommended that the EPA further evaluate the statistical power of the model and monitoring to distinguish between non-attainment rates of differing magnitude. With the information in hand, VAMSA concludes that the "1% rule" should be a "4% rule" at minimum.

4. *Inaccuracy of groundwater inputs*: The model handles groundwater inputs/loads in a very simplistic manner that is dissimilar to physical reality. Or as stated by STAC (2008a), "the model does not represent the full coupling of the groundwater to the surface water system on a regional scale." Considering that 50% of the total freshwater flow to the Bay is derived from groundwater (Bachman and others, 1998), this is a major model limitation and source of uncertainty for management scenarios.

5. *Lack of criteria for acceptance of model predictions*: Predictions of dissolved oxygen and chlorophyll-*a* in some segments are characterized by anomalies (e.g., counterintuitive trends with decreasing loads. EPA recognized many of the most obvious problems, and used poor model behavior as a justification for not using

[2] Batiuk, R. and Shenk, G., 2010. Technical Rationale for Documenting Attainment for 1% Non-attainment Dissolved Oxygen Criteria Values. Attachment C2 for State/District Co-Regulators June 14, 2010 Conference Call (Attachment B to this Appendix).

DO or chlorophyll-*a* attainment in many segment-seasons (e.g., Keisman, 2010a; Keisman 2010b). However, in most of these cases, the underlying cause(s) were not identified, and full implications of these problems for the model were not explored. The same problems that caused obviously poor model behavior in some segment-seasons might be also causing more widespread but less obvious problems in other segment-seasons. We see no evidence that the CBPO developed objective criteria for the acceptance or rejection of model results in these circumstances. Poor behavior of the James River chlorophyll-*a* model is discussed in the VAMSA comments at Section V.

6. *Poor chlorophyll-a calibration*: The chlorophyll-*a* calibration is obviously very poor in many segments (e.g., tidal freshwater James), and EPA has not demonstrated that the model is a useful predictor of annual changes in chlorophyll-*a* in other key segment-seasons. This comment is discussed in more detail in Section V.

7. *Instability and inaccuracy in urban land use assumptions*: The watershed model suffers from questions regarding accuracy of the urban land use acreages. Urban land use breakdowns have been very unstable between model versions and even subversions, varying with different derivation methods and assumptions. For example, the urban land use breakdown varied by millions of acres between model version 5.2 and 5.3. It is unclear whether the latest version is accurate or has been adequately ground-truthed. Urban stormwater loads and implementation costs are highly sensitive to the assumptions regarding urban land use breakdown.[3]

8. *Missing point sources*: VAMSA has learned from VAMWA that the current version of the model framework does not include 139 active Virginia point sources. Further, EPA is aware of this error, however it has not been corrected due to a lack of time until EPA's self-imposed December 31, 2010 deadline.

9. *Inappropriate application of watershed model to local level.* In their review of the Phase 5 watershed model, STAC (2008) clearly stated that the model was not appropriate for use at the local level, and would need recalibration/resegmentation for this application. It is unclear, then, why the Bay Program is continuing to promote the application of the model to determine local-level loads and allocations, and why EPA is calling for such values in the Phase 2 WIPs.

10. *Overparamterized modeling framework*: The model combined modeling framework is so complex and highly parameterized that there are no unique calibration solutions; it is easy to obtain the "right" answer for the "wrong" reason. Calibration also relies on regional calibration factors that act as "black box" knobs, divorcing the model result from physical understanding of the

[3] Materials at Attachment C to this Appendix.

processes. While necessary for calibration, these factors introduce yet another source of uncertainty into model predictions.

11. *Inconsistent watershed model results*: We understand that a consultant retained by another stakeholder has run the watershed model has obtained widely different results on different computers. We encourage the Bay Program to fully investigate the reasons and implications of this finding.

B. EPA's Critical Period Is Appropriate

VAMSA concurs with EPA's decision to use 1993-95 as the critical period for the nutrient TMDL.[4] This period had relatively high winter-spring inflows, but not so extreme that the TMDL would be based on an extremely rare hydrologic event. A TMDL based on 1993-95 hydrology will be protective under the great majority of hydrologic conditions.

C. EPA's Use of an Implicit Margin of Safety Is Appropriate

The Draft TMDL depends on a very complex framework of water quality standards, assessment methodologies, and models to derive allocations; each with its own environmental conservatism. This combined framework results in a sum level of conservatism reflecting all of the contributing sources of conservatism. For example, the water quality criteria themselves are conservative, as stated in the original criteria document (EPA CBPO, 2003):

> ...these criteria were developed with conservative (protective) assumptions, allowing a small percentage of circumstances in which the criteria may be exceeded will still fully protect the tidal-water designated uses.

The assessment methodology includes several conservative elements, such as the fact that any exceedance of the cumulative frequency distribution ("CFD") reference curve is considered a potential violation, even if the segment being assessed has a lower total violation rate in time-space (*i.e.*, area under the CFD curve) than the reference condition. The use of the default 10-percent reference curve for some criteria is also conservative in that Bay sites that are believed to be complying with standards are being found not to be in compliance based on conservative assumptions of the TMDL. The fact that the TMDL is developed for a critical 3-year condition, instead of average conditions, provides another layer of conservatism.

Furthermore, although the model is not designed to be explicitly conservative, a review of the UMD/MAWP Year 1 and Year 2 BMP efficiency reports revealed many examples of where conservatively low BMP efficiencies were selected for use with the Phase 5 watershed model. For example:

BMP	Conservative Assumption from Year 1 & 2 BMP Efficiency Reports
Riparian buffers	"...a 20% reduction in the effectiveness values is applied to

[4]<u>See</u> July 16, 2009 Technical Memorandum from C. Bell to C. Pomeroy (Analysis of January-May Inflows to the Chespeake Bay during the 1996-98 Period) and other materials (Attachment D to this Appendix).

	efficiencies from literature sources…"
Urban wet ponds and wetlands	"The uncertainty in how improper maintenance will adjust BMP efficiencies supports the recommendation to use a more conservative percent removal estimate."
Dry detention basins	"…effectiveness estimates for Dry Detention Ponds/Basins and Hydrodynamic Structures were not changed based on the recommendation of the USWG. However…the available literature does suggest somewhat higher removal rates…"
Bioretention	"The 10% TN concentration reduction [is] a conservative judgment…"
Vegetated open channel	"A more conservative value from the CWP estimate was selected…"
Permeable pavement	"…a conservative approach is taken to estimating permeable pavement and paver performance."
Infiltration basins and trenches	"…a 15% reduction in TN is used here for systems with sand or vegetation, and 0% TN removal for systems without sand and/or vegetation, to be consistent with the other infiltration and filtration BMPs in this report and to be conservative."
Off-stream watering	"…we proposed values close to the conservative literature base…"

The Bay Program Office has identified specific sources of environmental conservatism that are built into the analysis that justify an implicit margin of safety for the TMDL:

- The fact that allocations to achieve D.O. standards are driven by a relatively small area in the Bay (segment CB4), and that most of the rest of the Bay system would achieve DO standards under higher nutrient loading levels.

- The fact that 100% of point sources are assumed in model scenarios to operate at their maximum permissible loading levels, which is highly unlikely to ever occur.

Given the multiple layers of conservatism in the TMDL allocation process, VAMSA supports EPA's decision to use an implicit margin of safety.

D. EPA's Failure to Recognize Essential Equivalency in Its Target Load Options is Unreasonable

In the determination of basin nutrient loadings (190 TN and 12.7 TP) EPA utilized the 1% rule to determine compliance (with the exception of certain problem segments). VAMWA's consultant, Clifton Bell of Malcolm Pirnie, performed a statistical "power analysis" to evaluate the minimum difference in D.O. that would be statistically detectable in the Chesapeake Bay Monitoring Program.[5] Based on the results of this analysis, segments that are close to attainment would require spatial D.O. violation rates that differ by 4% or more before they would be

[5] See Attachment E to this Appendix.

statistically distinguished from one another. The management implications are that Bay model D.O. scenario results with differences less than 4% should be considered "essentially equivalent." This is not the case in the current TMDL. Based on the above referenced "power analysis," the scenario associated with Target load Option A produces results that are "essentially equivalent" to EPA's recommended basin target loads of 190 mpy/yr TN and 12.7 mpy/yr TP. At this level of nutrient loading the key Bay segments of CB4MH, CB5MH, MD5MH, and VA5MH are predicted to be in attainment or be within 2% of attainment. It is recognized that Target load Option A would not immediately address attainment in some of the side segments. However, effectively addressing these side segments would require separate, locally oriented modeling analysis with tools better adapted to evaluating local conditions. The Target Load Option A to comply with D.O. standards in the main bay is essentially equivalent to the more stringent and costly to attain allocations associated with 190 TN and 12.7 TP and the TMDL; this must be recognized in the TMDL.

E. EPA Should Assume Better Design, Installation, Operation and Maintenance for Modeled BMPs

It is well known that historically many non-point BMPs have not been accompanied by programs or methods to ensure proper design, installation, operation, or maintenance. It is reasonable that model calibration scenarios should assume, at a minimum, historical "average" management conditions. Any other approach—including the use of conservatively low values—would make the model less accurate and force management decisions that may be more costly and/or provide less benefit. However, it is not necessary for forward-looking management scenarios to retain the assumption of historically-average BMP management. Rather, improvements in the way BMPs are installed, operated, and maintained are a viable implementation component. Modeled TMDL allocations scenarios should reflect the manner in which BMPs *should* be designed, operated, and maintained, not necessarily how they have historically been managed.

One example of where EPA and the Bay States have assumed a high level of nutrient removal performance is for wastewater treatment plants. The performance expected and used in the model is based on properly installed, operated and maintained facilities. The standard for performance relative to design of any nutrient removal strategy (wastewater plants, BMPs, filter feeders, etc.) used in the Bay model should not be different.[6]

These actions would improve the effectiveness of BMPs to reduce loads and improve reasonable assurance of reductions from these sectors.

[6] See VAMWA Chesapeake Bay Team Memo re BMP Efficiencies to VAMWA and MAMWA Boards of Directors, January 21, 2009 (Attachment F to this Appendix).

ATTACHMENT A

Scientific and Technical Advisory Committee

Chesapeake Bay Watershed Model Phase V Review

February 20, 2008

Lawrence Band[1], Theo Dillaha[2], Christopher Duffy[3], Kenneth Reckhow[4], Claire Welty[5]

[1]University of North Carolina Chapel Hill, [2]Virginia Tech, [3]Pennsylvania State University, [4]Duke University, [5]University of Maryland Baltimore County

STAC Publication 08-003

About the Scientific and Technical Advisory Committee

The Scientific and Technical Advisory Committee (STAC) provides scientific and technical guidance to the Chesapeake Bay Program on measures to restore and protect the Chesapeake Bay. As an advisory committee, STAC reports periodically to the Implementation Committee and annually to the Executive Council. Since it's creation in December 1984, STAC has worked to enhance scientific communication and outreach throughout the Chesapeake Bay watershed and beyond. STAC provides scientific and technical advice in various ways, including (1) technical reports and papers, (2) discussion groups, (3) assistance in organizing merit reviews of CBP programs and projects, (4) technical conferences and workshops, and (5) service by STAC members on CBP subcommittees and workgroups. In addition, STAC has the mechanisms in place that will allow STAC to hold meetings, workshops, and reviews in rapid response to CBP subcommittee and workgroup requests for scientific and technical input. This will allow STAC to provide the CBP subcommittees and workgroups with information and support needed as specific issues arise while working towards meeting the goals outlined in the *Chesapeake 2000* agreement. STAC also acts proactively to bring the most recent scientific information to the Bay Program and its partners. For additional information about STAC, please visit the STAC website at www.chesapeake.org/stac.

Publication Date:
February 2008

Publication Number:
08-003

Cover photo of the Sassafras River provided by Jane Thomas, Integration and Application Network (http://ian.umces.edu/).

STAC Administrative Support Provided by:
Chesapeake Research Consortium, Inc.
645 Contees Wharf Road
Edgewater, MD 21037
Telephone: 410-798-1283; 301-261-4500
Fax: 410-798-0816
http://www.chesapeake.org

Chesapeake Bay Watershed Model Phase 5 Review
Review Panel: L. Band, T. Dillaha, C. Duffy, K. Reckhow, C. Welty
February 20, 2008

Introduction

In the fall of 2007 the Scientific and Technical Advisory Committee (STAC) of the Chesapeake Bay Program (CBP) recruited the authors as an independent panel of experts to review the Chesapeake Bay Watershed Model (CBWM) Phase 5 effort and make recommendations for its enhancement. The review panel met as a group on January 23 - 25 in Annapolis, MD. Limited documentation on the Phase 5 CBWM was provided in advance. Presentations were given to the review team by Richard Batiuk, Gary Shenk, and Lewis Linker of the EPA Chesapeake Bay Program. Many of the documents distributed for review prior to the meeting in Annapolis were in draft form, with key components missing or incomplete. On the first day of the review, CBP personnel presentations provided a more detailed description of the Phase 5 model components and calibration process and an update on the status of the Phase 5 model calibration and validation, which were in progress. On day two of the review, CBP personnel responded to additional panel questions and the panel began to conduct the formal review. This document summarizes the panel's assessment of (1) work to date, (2) the model's suitability for making management decisions at the Bay Watershed and local scales, and (3) potential enhancements to improve the predictive ability of the next generation of the CBWM. The reader should be aware that model documentation required for this review was incomplete and this review is based solely on the information provided. Improved and continuous documentation of the model and data environment should be implemented as soon as possible.

The CBP represents one of the largest and most complex watershed management efforts in the U.S. and its success is partially contingent upon the accuracy of the CBWM. The task demands a detailed description of hydrological, biogeochemical and climatological processes over a multi-jurisdictional regional watershed scale. Data demands are daunting and differentially available over the watershed. While more process-oriented research models are available, they are not yet feasible for the geographical scale of the CBP, and currently do not have the ability to simulate all the complexities of the Chesapeake Bay Watershed (CBW) required for CBP management decisions. The CBWM represents a significant simplification of the CBW with significant compromises; however, we believe that the CBWM is appropriate given the scale, complexity and mechanistic basis of the modeling and management frameworks that are feasible with the current state-of-the-science of watershed modeling for management purposes. The envisioned Chesapeake Bay Program Community Watershed Modeling effort is promising and provides the potential to engage a much larger community in the development and application of the CBWM.

Additional complexities that the next generation of the CBWM should address include: (1) accounting for the fact that much of sediment and nutrient transport into the Bay may take place during annual extreme events (these large events are responsible for much of the excessive erosion and flushing of stored materials as well as CSOs and SSOs (combined and sanitary sewer overflows)); and (2) the fact that management can involve significant time lags in terms of the timing between management changes and subsequent environmental response. We are concerned that the present CBWM may not be capturing these complexities adequately.

The CBWM modeling team has done an extraordinary job of pulling together the information base from disparate sources, designing and implementing a set of software tools and methods to integrate a data and modeling system. This has been done with extremely limited personnel and resources (monitoring, programming, disciplinary expertise, etc.).

It is important to note that the Phase 5 Watershed Model is not a strict implementation of Hydrologic Simulation Program-Fortran (HSPF) as was implemented in previous versions of the CBWM. The Phase 5 model is a melding of two major components of HSPF, the land segment and reach simulation modules, with the External Transfer Module (ETM), which modifies output from the HSPF land segments to account for the effects of the presence or absence of BMPs on sediment and nutrient loading to reaches. The Phase 5 model also includes interfaces with other models such as the airshed, estuarine, and land use change models and various other modules, which create the required UCI (input) files for the land segment and reach models.

As in HSPF, the CBWM is a lumped, conceptual representation of the watershed. The conceptual stores and fluxes, which are lumped at the subwatershed level (subwatersheds average 66 mi^2 in size), make it difficult to relate measured quantities such as soil moisture, groundwater levels, and soil and sediment chemistry to simulated values. The choice of the subwatershed level is a critical scale choice as the model maintains a one-to-one mapping of stream or river reach to contributing subwatershed area. For application to the full CBW, the Phase 5 CBWM uses a threshold scale of representing the extent of the river network and corresponding subwatershed partition to streams with at least 100 cfs mean annual flow (or 50 cfs if the subwatershed is gauged). This has the effect of eliminating smaller streams and their channel processes, and their effects are implicitly lumped or included in terrestrial processes. The scale choice is based on data availability, available resources (e.g. personnel, budget) to parameterize the model, and computational limitations.

The current implementation of the model is mainly geared towards the scale of the major tributaries and the Chesapeake Bay Basin. A project of this scale requires a modeling and information environment to formalize the approach within a systems framework. This framework is still evolving.

The review panel was asked to address the following four questions:

1. Are the model structure, dynamics, and calibration sufficient for the management purposes at the regional scale to support Chesapeake [Watershed] water quality management with regard to segmentation, land uses, HSPF modifications, and ancillary software?

2. Are the model structure, dynamics and calibration sufficient for the management purposes at the local watershed scale to support sediment and nutrient TMDLs with regard to segmentation, land uses, HSPF modifications, and ancillary software?

3. Are the data inputs sufficient to support management decisions with regard to meteorology, nutrient inputs, land use, BMPs, septic systems, point sources, and atmospheric deposition at the regional and local scales?

4. Phase 5 is the latest generation of a model that has been applied in the Chesapeake watershed for more than two decades. To address increasingly complex and local-scale management needs anticipated in the watershed, what should the next generation of the Chesapeake Bay Community Watershed Model look like?

These four questions address the utility of the model for management purposes at both the regional (major watershed tributaries) and local (~66 mi^2 subwatershed) scales.

Response to Specific Questions

1. *Are the model structure, dynamics, and calibration sufficient for the management purposes at the regional scale to support Chesapeake [Watershed] water quality management with regard to segmentation, land uses, HSPF modifications, and ancillary software?*

   a. Before this question can be answered fully, model calibration and validation must be completed, documented and re-reviewed since the panel only had the opportunity to review draft model documentation and to evaluate preliminary calibration and validation results. While a substantial number of model simulations have been produced and compared with time series of flow, and sediment and nutrient concentrations and loads, this information must be summarized at the scale of the major CBW tributaries. The calibration strategy appears to be innovative and sound, but it is difficult to judge until completed. The time series comparisons that were presented to the review team were interesting, but did not convince the panel that an adequate calibration had yet been achieved beyond streamflow. Although Question 1 does not directly address validation, we feel that validation is essential and a required step in model development, particularly if the model is to be used for TMDL development purposes. The current validation strategy -- selecting validation time periods within the calibration period-- is not a good one, as this is likely to simply re-confirm the results from calibration periods that are adjacent-in-time to validation periods (which might result in the validation period being essentially equivalent to the calibration period). A much better strategy is to completely separate calibration and validation time periods - for example, calibrate with the 1985-95 data and then validate with the1995-2005 data. If the results of the validation exercise suggest that the calibrated model is flawed, then the validation results can be used to reformulate the model. In that case, the best option for re-validation would be to use the original calibration data set for validation of the revised model.

   b. We still believe that uncertainty analysis is essential. We understand that the model is very consumptive of computer time to operate for the full CBW. However, uncertainty analysis could provide the basis for the "margin of safety" (MOS) used in the TMDL plans. We see two options for this difficult problem: (1) use the difference between predictions and observations during the validation period to serve as a measure of prediction uncertainty, or (2) following the 2005 review recommendation, use one or two of the tributaries, or representative subwatersheds of a tributary, for this purpose. This would reduce the amount of computer time necessary to run multiple realizations.

c. We have concerns regarding the representation of BMPs in the model. Several BMPs (improved nutrient management and low till row crops) are implemented as separate land uses reflecting altered management and appear reasonable. Other BMPs are simulated as edge-of-field (EOF) or edge-of-stream (EOS) practices and their effects are simulated using constant (0-1) efficiency factors drawn from the literature and best professional judgment. There are two specific concerns with this approach:

(1) In many cases, these latter BMPs may not conserve mass. Removal of sediment and nutrients are not explicitly accounted for in the model mass balance. A means must be found to account for and simulate the long-term fate of sediment and nutrients that are "trapped or removed" by BMPs if they are not permanently removed (e.g., denitrification or transport out of the watershed). As an example, build-up of sediment or nutrients in a buffer or wetland may lead to reduction in removal efficiency over time or conversion of the BMP to a source under certain conditions. At present, no build-up of mass in these BMPs is simulated, nor is subsequent release during extreme events permitted.

(2) Removal efficiencies of BMPs are known to be dependent on climate, flow rates, hydrogeologic setting, and implementation and maintenance conditions. Within the External Transfer Module (ETM) framework, these efficiencies are currently fixed at constant values. However, they could either be sampled from a distribution function (with form and bounds set from the literature) or conditioned on flow rates (if appropriate). This would allow "breakthrough" of sediment and nutrients for a subset of the population of BMPs, which could have important downstream impacts.

d. The limited coupling of the land segment and river reach modules does not allow for overbank deposition, or other important loss rates from the river reach system under high flow conditions or under extreme drought (if we understand the model correctly). This may bias total export predictions but we note that a much more detailed model would be required to address these issues. A similar situation exists for dynamic interactions between wetlands and stream reaches. These issues should be dealt with in the next generation of the model.

e. The model currently is implemented with a representation of river reaches with mean annual flow exceeding 100 cfs (or 50 cfs for gauged watersheds), which fails to account for smaller streams and the heterogeneity of small watersheds that can influence BMP performance and the development of management options and TMDLs.

f. Validation has been conducted by choosing specific years within the 1985-2005 domain to use as validation periods. This approach does not account for long-term changes and the stability of the model parameters over a period that may have significant change in climate, land use or management options. Instead, we recommend that the modeling team identify those watersheds with sufficient hydrologic, nutrient and sediment records to allow an initial calibration period (e.g. 1985-2000), and a subsequent contiguous validation period (e.g. 2001-2005). These periods may vary in length and time for the different stations depending on the availability of data. It is not necessary or feasible to validate each watershed given current data, but additional monitoring and use of other existing data sources not currently being used should be used to evaluate model performance in key subwatersheds in the Phase 5 modeling effort. Validation efforts should focus on those watersheds with adequate observed data for calibration and validation.

g. The model does not represent the full coupling of the groundwater to the surface water system on a regional scale. It is believed that a significant percentage of nitrate load to the Bay is from direct groundwater inputs. Our understanding is that this is not fully captured by the model. A means should be found to capture this load if it is significant for management decisions if possible. Otherwise this should be given as a model limitation.

h. The model does not capture long-term persistence such as drought flows because of lack of coupling between surface water and groundwater. This deficiency also affects nutrient loads as mentioned above.

2. *Are the model structure, dynamics and calibration sufficient for the management purposes at the local watershed scale to support sediment and nutrient TMDLs with regard to segmentation, land uses, HSPF modifications, and ancillary software?*

We define the "local watershed scale" as the current lowest level of CBWM segmentation, characterized by reaches with mean annual flow > 100 cfs (~66 mi^2 area on average).

a. This question was discussed at length with the CBWM team. We agree with the team that the current CBWM implementation is not appropriate for development and implementation of TMDLs at the local watershed scale. A major barrier appears to be the scale of information built into the CBWM, which is based on the county level data and river reach segmentation at the 100 cfs threshold and designed for full watershed or major tributary scale analysis.

b. A potential approach is to make use of community modeling framework in which local watershed managers could make use of additional modeling tools and data to resegment, recalibrate and implement the model at appropriate local scales using more site specific local information. Local-scale data can be obtained from specific sampling and measurement, or from higher-resolution spatial data sources and modeling tools.

3. *Are the data inputs sufficient to support management decisions with regard to meteorology, nutrient inputs, land use, BMPs, septic systems, point sources, and atmospheric deposition at the regional and local scales?*

<u>Response for Regional Scale</u>

a. Yes, with the following qualifications. We assume regional scale to mean major watersheds e.g., from the scale of the Patuxent to the Susquehanna River Basins. The data on meteorology, land use, point sources, and atmospheric deposition appear to be of sufficient quality at this scale. At the county level there appears to be reasonable estimates of fertilizer sales, which are used to estimate nutrient inputs at the county/subwatershed scale. Data available from soil-testing laboratories/programs could possibly be used to estimate soil phosphorous pools for the models. BMP efficiencies are estimated from literature values, expert judgment, and county-level data bases. BMPs are being represented in the simplest way possible (described previously); representation of BMPs statistically and dynamically is important. In terms of annual changes, this can be represented by the model (data on BMPs can be changed annually). As in the model review recommendations of 2005, we recommend/encourage the modeling team to compile account for the dynamic behavior of BMPs with respect to their efficiencies.

b. We are concerned about the low-order meteorological interpolation as it has the potential to oversmooth weather patterns, leading to a loss of information about local extremes. The inaccuracies of precipitation timing will significantly affect the hydrology modeling. We recommend considering use of the bias-corrected and merged NEXRAD-gauge precipitation data (1 km^2 grid) as it becomes available, and to evaluate the current precipitation product for use prior to the period of NEXRAD availability.

<u>Response for Local Scale</u>

We believe that it is inappropriate to use the existing CBWM county and subwatershed data sets for local-scale modeling applications. Data must be disaggregated at a finer scale for local scale applications.

In addition to the national 30-m data sets for land cover and soil surveys, there are a number of small-scale watersheds (< 100 cfs) within the CBW that have fine-scale temporal and spatial data sets available (e.g. weekly chemistry, LiDAR, more detailed land cover and infrastructure, etc.) that can be used for smaller-scale modeling applications. Examples include the Baltimore Ecosystem Study Long Term Ecological Research site; SERC research sites; the Penn State Critical Zone Observatory (Susquehanna/Shale Hills/Leading Ridge); Virginia's Nomini Creek, Owl Run, Polecat Creek, Long Glade and Mossy Run watershed studies; and USGS and ARS research sites and watershed monitoring studies.

4. *Phase 5 is the latest generation of a model that's been applied in the Chesapeake watershed for more than two decades. To address increasingly complex and local-scale management needs anticipated in the watershed, what should the next generation of the Chesapeake Bay Community Watershed Model look like?*

Our comments below address the CBWM and do not address the Chesapeake Bay Community Watershed Modeling effort as it is not currently operational.

a. Long-term mass balances. The Chesapeake Bay restoration and other large-scale watershed and ecosystem projects are addressing processes and management actions that occur and will have impact over decades. Over this period of time, intentional and unintentional changes in the characteristics of the watersheds will occur, including land cover, climate change, land management, and ecological succession. Over short time scales these may be prescribed, whereas over long time scales allowance has to be made for interactions and feedbacks among these processes. As an example, in the current model, mass is not fully conserved in the methods used to simulate BMPs and deep groundwater percolation. Groundwater flows, BMPs, and other processes should be changed so that mass balance is maintained.

b. Process-oriented, distributed modeling at the sub-basin scale. The CBWM is derived from an older paradigm that was not designed to produce state or flux variables that can be easily measured, except for stream flow (e.g., soil moisture and tension, groundwater levels, water vapor flux). The model would be more useful if there was an ability to compare a greater number of measured and modeled variables over space and time. This could include such variables as rooting zone soil moisture and groundwater depths.

c. Distributed approach. We recommend moving from a lumped conceptual model at the subwatershed scale to a more distributed parameter approach that simulates processes at smaller scales. We have the ability to make many more measurements now than we did at the time HSPF was formulated, both across different variables and at different scales. Therefore any new model development should take advantage of new measurement technologies (e.g., ADCP, satellite data (e.g., canopy LAI, productivity, surface temperature), sap flux, LiDAR, high resolution aerial photography, eddy covariance stations, continuous real-time nutrient and chemical sensors, sensor network technologies, and isotope lasers) to improve the temporal and spatial resolution of model inputs.

d. Ecosystem dynamics. The next generation model should incorporate a dynamic ecosystem approach that integrates and fully couples carbon and nutrients in the soil and water cycles and incorporates spatially explicit land management activities.

e. Parallel computer processing. The next generation CBWM should be designed to take advantage of the capabilities of parallel computing to allow watershed coupling and feedback, reduce computational requirements, and facilitate analysis of integrated management alternatives.

Suggested Implementation Time-Line and Additional Recommendations

The following actions are suggested to improve the use of the CBWM for management and TMDL development purposes.

Immediate Needs

1. A much higher level of resources is needed for adequate model development, calibration, and validation. It is remarkable what has been accomplished, but the effort is too dependent on too few highly-trained personnel. Given the great importance of this effort to the success of the Bay in terms of achieving water quality goals, the modeling effort appears to be grossly underfunded. A reasonable approach is to implement a working design team of CB plus outside scientists and engineers with technical support to begin the design and testing of new and existing models that specifically deal with these questions. The effort is critical to the success of the Bay program and achieving the Bay TMDL. A modeling budget double or triple the current level of funding for

the next two to three years will likely be required for the development of Chesapeake Bay TMDLs that can withstand court challenge.

2. The model documentation, calibration, and validation must be completed so that these items can be reviewed by the scientific and user community. The model documentation should be continuously updated. The calibration efforts should be documented on subwatersheds and watersheds with adequate monitoring data. Validation efforts should be limited to subwatersheds and watersheds with adequate monitoring data.

3. There should be an increase in, and cross training of, modeling team members so that modeling efforts are not dependent of the skills and knowledge or loss of single team members. The team has expertise in hydrology/water quality modeling. It needs additional expertise in computer programming, agricultural nonpoint source pollution control, urban nonpoint source pollution control, TMDL development, groundwater hydrology/modeling, instream processes, etc. Additional personnel do not necessarily have to be full time, but they must be engaged with the effort and be able to work with the CBWM team on a regular (weekly) basis.

4. The monitoring to support CBWM development, calibration, and validation should be improved. In terms of monitoring, given the investment in the 20-year history of the modeling program and the envisioned costs of restoration, it is remarkable that there are only three continuous daily nutrient and sediment monitoring stations (our understanding) in the entire 64,000 sq mi Chesapeake Bay watershed. Given the advancements in sensor and sensor network technology, it is of paramount importance to invest in this technology and link it to the modeling effort to improve the model calibration quality. The monitoring could also be tied to the intensive subwatersheds mentioned in (2) above.

5. We were very impressed by the creative methods used to automate and improve calibration by focusing on specific properties of the streamflow time series and relationships among model parameters. We recommend that this approach be explored further.

6. Although major changes have been undertaken to develop the current model, major software engineering needs to be undertaken to streamline the code, make input and output processing more efficient, and utilize interactive web-based visualization software. The Chesapeake Bay Community Modeling Program has started to do this, although this is not yet operational.

7. Calibration and validation could be improved by using a variety of additional tools: temporal aggregation, disaggregation (Bo, Islam, Eltahair, 1994, Water Resources Res., 30(12), p. 3423–3435, smoothing, and space-time principal components analysis (Elsner and Tsonis, 1996, Singular Spectrum Analysis, Springer, 177pp). A good effort in this area has been made in the innovative calibration methods that seek to preserve important properties of the hydrograph, e.g., recession rates.

8. Uncertainty analysis. There is a need to develop some uncertainty measure on predictions. One possibility is to develop a standard error calculation based on predicted versus observed values during validation; this could be the basis for the margin of safety (MOS) calculations needed for TMDLs. For longer time series of available data, recalibration of the model could be

used to evaluate the stability of parameters as a function of time to determine whether they are stable or drifting.

9. There should be a more cleanly thought-out scenario process. We understand that the scenario development is not fully controlled by the modeling team, but there may be some schemes developed to categorize and catalog different types of scenarios so that a master database of model responses to different management scenarios is available without running the model. This can be used both to aid managers who may be able to base planning on previous results, identify missing key scenarios, or serve as a basis for a data mining approach to formulate simpler models or emergent properties or behaviors of the Chesapeake Bay Watershed.

10. An assessment should be made of the use of county-level data from state soil testing labs to set initial soil nutrient level pools of major soils, crops, and land uses and update pool concentrations over time if soil testing lab data indicates changes. The approach used to quantify soil nutrient pools and fluxes should be changed so that nutrient pools are not calibrated.

11. New land uses should be added so that appropriate BMPs can be simulated using HSPF itself (as with low till cropland and improved nutrient management) rather than BMP efficiency factors.

12. Procedures should be developed to simulate the dynamic nature of BMPs and the sensitivity of BMPs to extreme events.

13. It is important to continue the development of a Chesapeake Bay Program Geodatabase as has been discussed at STAC and CCMP meetings. This standardizes all data within the Bay and Watershed and allows wider use and application through standardization.

<u>Intermediate Needs (1 to 3 years)</u>

1. The model should be used to identify subwatersheds that deliver disproportionate sediment and nutrient loadings to the Bay and that have disproportionate impacts on Bay water quality during critical periods. This could be used to target Bay implementation activities to the most cost effective sources.

2. There should be an applied research program established by the CBP to improve our understanding and ability to model key processes affecting sediment and nutrient transport in the CBW. The research program should be directed towards achieving the science and management goals of the watershed component of the Bay program.

3. Improved representation of channel erosion, scour and deposition dynamics is needed. The possible use of components from the CONCEPTs or other channel erosion models should be investigated.

4. Action should be taken to proactively identify and consider future threats to future water quality (e.g., thermal waste heat from power generation, ethanol waste fertilizer issue, dredge spoil disposal, allocation issues) and identify potential ways that they can be simulated in the

model should the need arise. This may be an appropriate activity for the applied research program.

Long-Term Needs (4 to 6 years)

1. Adequate funding and resources must be provided for an integrated modeling and monitoring program to enhance modeling efforts.

2. A new generation of the CBWM is needed that is:
   a. Not based on HSPF
   b. Process-oriented and represents
      - Instream processes (interactions between biotic and abiotic components of the ecosystem)
      - Dynamics of BMPs – simulates BMPs through their effects on model parameters rather than with current efficiency factors and accounts for ultimate fate of "trapped" sediment and nutrients.
      - Evapotranspiration, crop growth, soil nutrient and carbon dynamics (continuous mass balance)
      - Groundwater dynamics, nutrient transport, and groundwater loadings to streams and directly to the Chesapeake Bay
      - Flood plain dynamics (interactions between sediments and nutrients in the flood plan and channels)
      - Wetland dynamics (interactions between wetlands and channel systems)
      - Priority pollutants other than sediment band nutrients

   c. A distributed parameter model
      - with much finer land segmentation and stream network representation
      - that is able to identify areas at the scale of 10 hectares that are disproportionately responsible for water quality impacts
      - that utilizes remote sensing data to estimate both historical and real-time model parameters

3. Potential to develop TMDLs for sediment and nutrients at the "local" scale.

Final Thoughts

Similar to the Everglades restoration in approach and complexity, the Chesapeake Bay restoration is dependent on a combination of integrated modeling, monitoring and expert judgment to forecast and guide management efforts with particular emphasis on nutrient and sediment management. Both efforts must develop and justify an integrated framework including the cooperation of multiple federal, state, local, public and private stakeholders in the design and implementation of a range of practices designed to reverse a large-scale eutrophication process. Management changes have a long-term memory. Persistence comes in over much longer time tables. The efforts will put in place strategies to alter hydrologic, ecosystem and social systems with the aim of preserving and improving valuable ecosystem services provided by the CB and the Everglades, understanding that there may be long term lags and feedbacks between the installation of the practice and significant outcomes.

Consequently, the restoration efforts in the Bay may yield much of the ecosystem services benefits of land management over a much longer term owing to time lag. It is essential that the Watershed Model, in conjunction with the linked atmospheric and bay models be able to represent these lags and feedbacks. In the Everglades, this has been approached by coupling a full ecosystem model with a distributed hydrologic simulation. A similar goal should be set for the CBW. In both cases of the CBW and the Everglades, the ability to develop and apply these models requires a significant amount of interdisciplinary data and observations to calibrate, verify, and guide model efforts. This should be a goal of the scientific and management communities.

ATTACHMENT B

Technical Rationale for Documenting Attainment for 1% Non-attainment Dissolved Oxygen Criteria Values

State/District Co-Regulators
June 14, 2010 Conference Call

Attachment C2

Rich Batiuk and Gary Shenk

Technical Rationale for 1%

- **Two separate analyses**
  - Evaluation of evidence for 'residual' of 1.5 percent or less non-attainment in dissolved oxygen criteria values across large ranges of model simulated load reductions across multiple Bay segments and designated uses (Batiuk)
  - Analysis of changes in the sensitivity of dissolved oxygen criteria attainment to simulated load reductions (Shenk)

Residual of 1.5% or Less Analysis

- Keep in mind:
  - 'Stoplight plots' already account for unallowable exceedences
  - Stoplight plots also account for restoration variances adopted in states' WQS regulations
  - Stoplight plots are the result of a comprehensive analysis system: model scenario output -> regression generation -> transformation of monitoring data -> criteria assessment fully consistent with states/DC WQS regulations

Residual of 1.5% or Less Analysis

- 21 designated use-segments with non-attainment values ranging from 0.0% to 1.5% (will round down to 1%) based on May 24, 2010 stoplight plot presented to WQGIT
  - Mainstem Bay, major river, small tributaries and embayment segments all represented
  - 11 open-water, 8 deep-water and 2 deep-channel designated uses

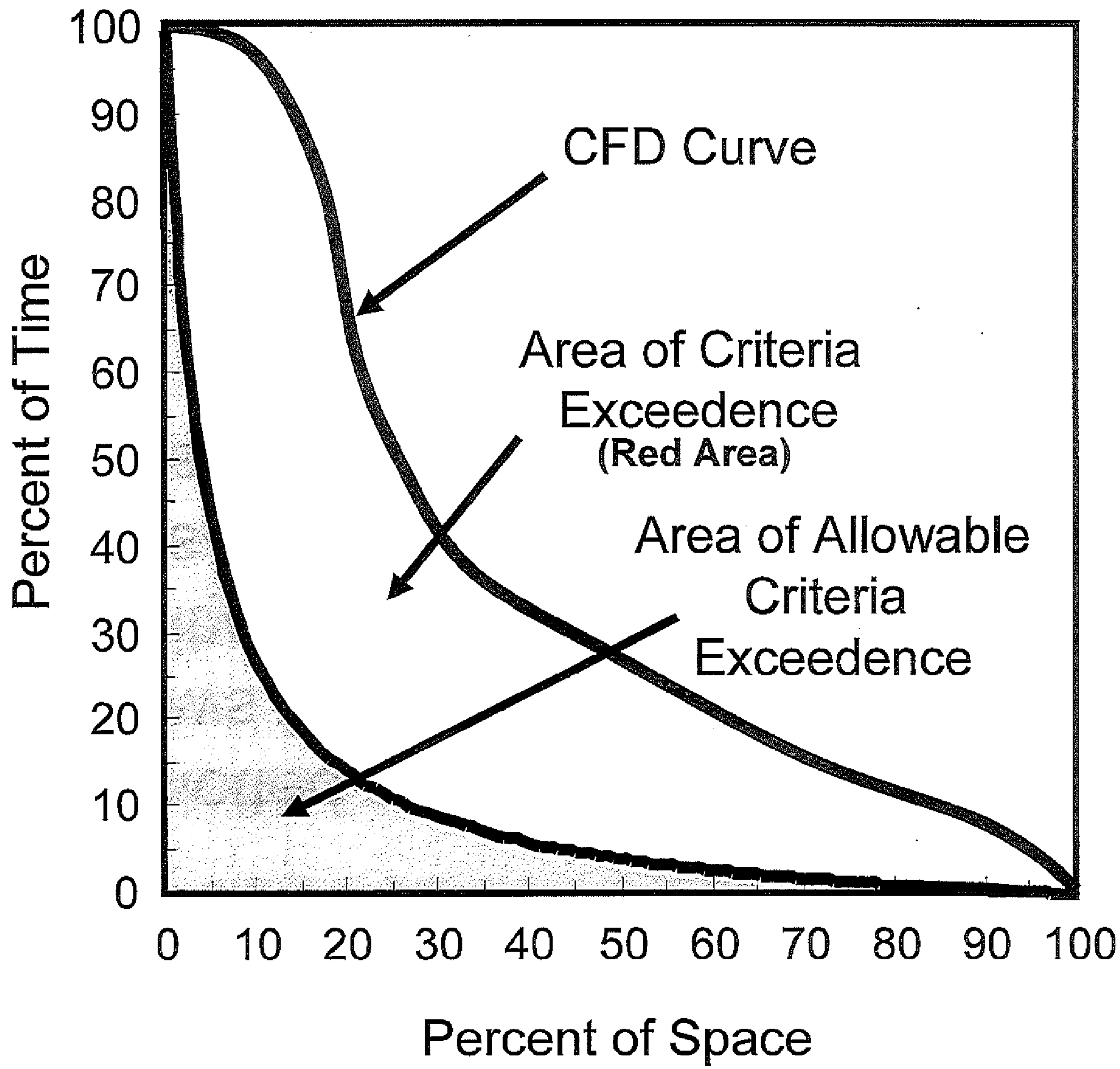
- Model simulated nitrogen load reductions ranged from 9 to 151 million pounds

Chesapeake Bay Segment	Designated Use	Criteria Non-attainment Range (%)	Model Simulated Nitrogen Load Range (million pounds/yr)
CB7	Open-water	0.5-0.0	200-141
CHOMH1	Open-water	0.1-0.0	254-179
CSHMH	Open-water	0.8-0.1	342-309
DCATF	Open-water	1.2-0.1	191-179
PAXTF	Open-water	1.0-0.6	190-179
DCPTF	Open-water	0.6-0.2	309-254
MAGMH	Open-water	1.3-0.3	342-191
MOBPH	Open-water	1.0-0.0	342-200
PIAMH	Open-water	0.1-0.1	191-179
TANMH	Open-water	1.5-0.1	342-309
YRKMH	Open-water	1.0-0.4	191-170
CB3MH	Deep-water	0.6-0.0	254-179
CB5MH	Deep-water	1.5-0.0	254-141
CHSMH	Deep-water	0.5-0.4	170-141
EASMH	Deep-water	0.8-0.2	200-170
MD5MH	Deep-water	1.5-0.1	191-141
MAGMH	Deep-water	0.5-0.5	170-141
PATMH	Deep-water	1.1-0.1	200-190
VA5MH	Deep-water	0.7-0.0	254-179
CB3MH	Deep-channel	0.2-0.1	200-190
EASMH	Deep-channel	1.3-0.0	190-170

Sensitivity to Load Reductions

- Plotting the change in unallowable exceedence (red area) per loading unit against the starting red area
  - Change in red area between two scenarios is divided by the change in the load
  - Changes in N and P loads are combined into a single measure:

Load units = (N + 10*P)/2


CFD Curve
Area of Criteria Exceedence
(Red Area)
Area of Allowable Criteria Exceedence
Percent of Time
Percent of Space
0
10
20
30
40
50
60
70
80
90
100

Sensitivity to Load Reductions

- When plotted against the starting red area, this allows for a direct comparison of sensitivity of the analysis system to load changes across different level of non-attainment

Sensitivity to Load Reductions

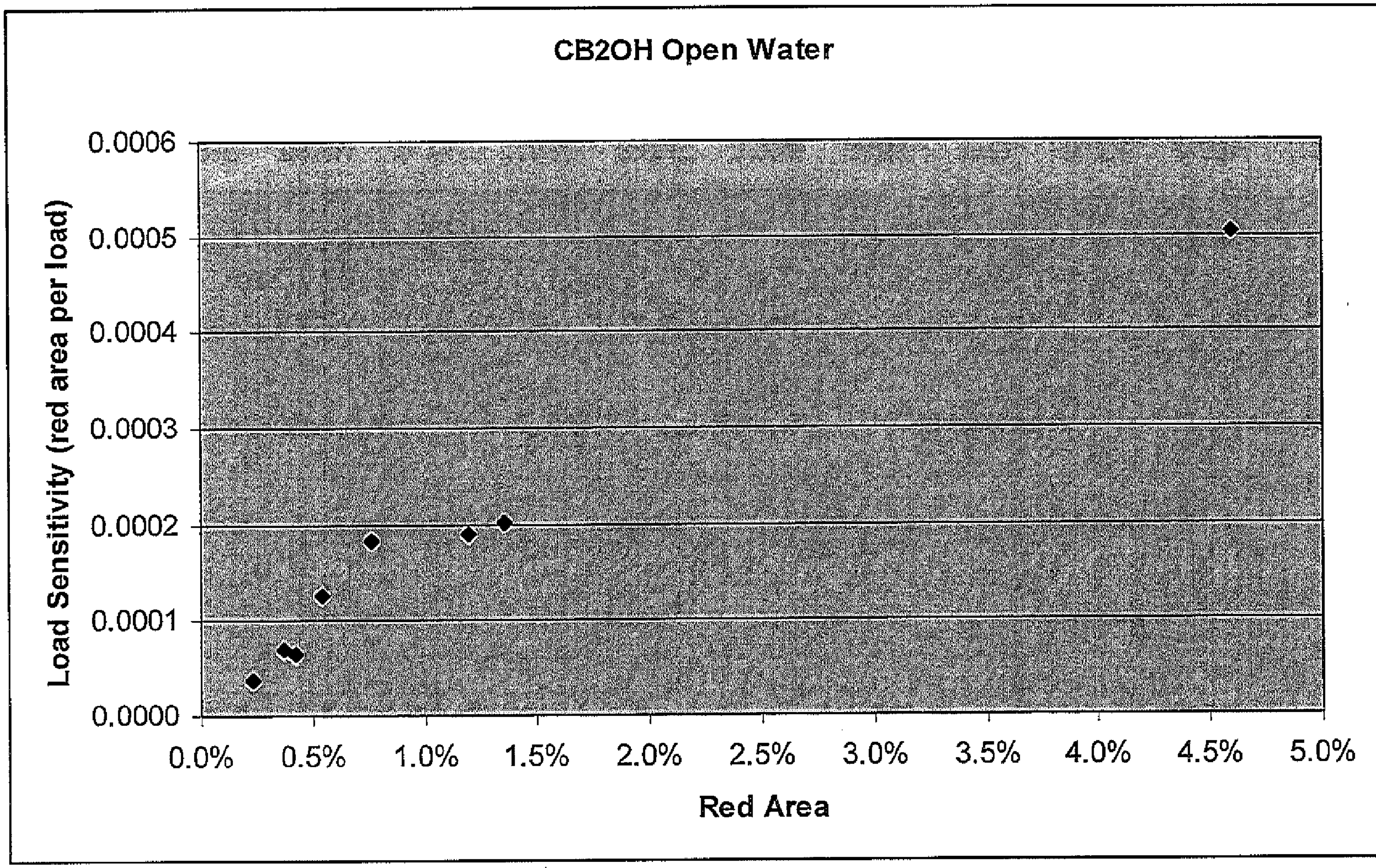
- 12 scenarios with eight 3-year periods for a total of 88 possible sensitivity assessments per designated use segment
  - Calculation involving scenarios where criteria were attained were not included in this analysis

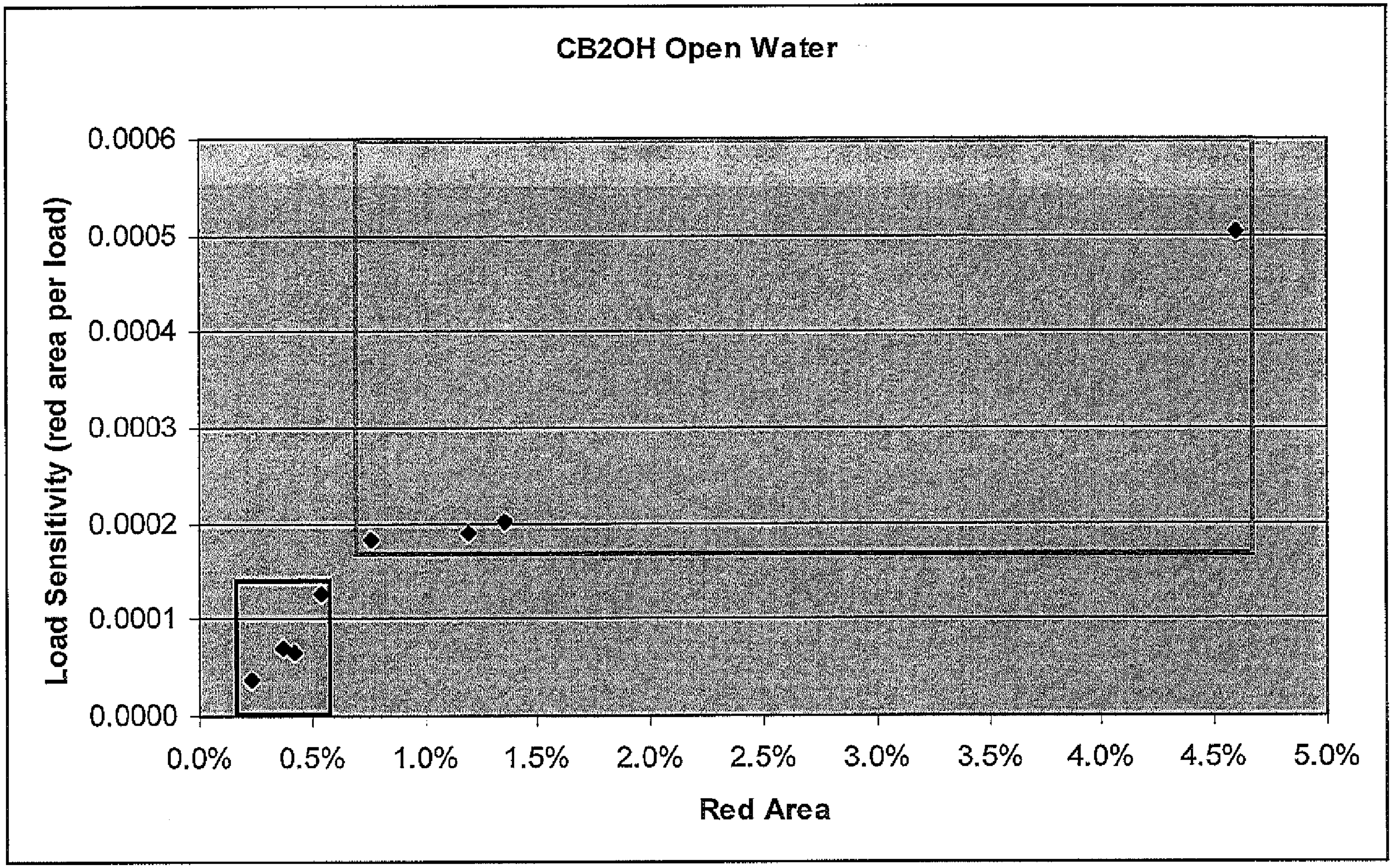
- Not amenable to tidal tributary segments
  - Loading are baywide, not specific to tributaries
  - Existing scenarios used for analysis have varying levels of reduction between different tributaries

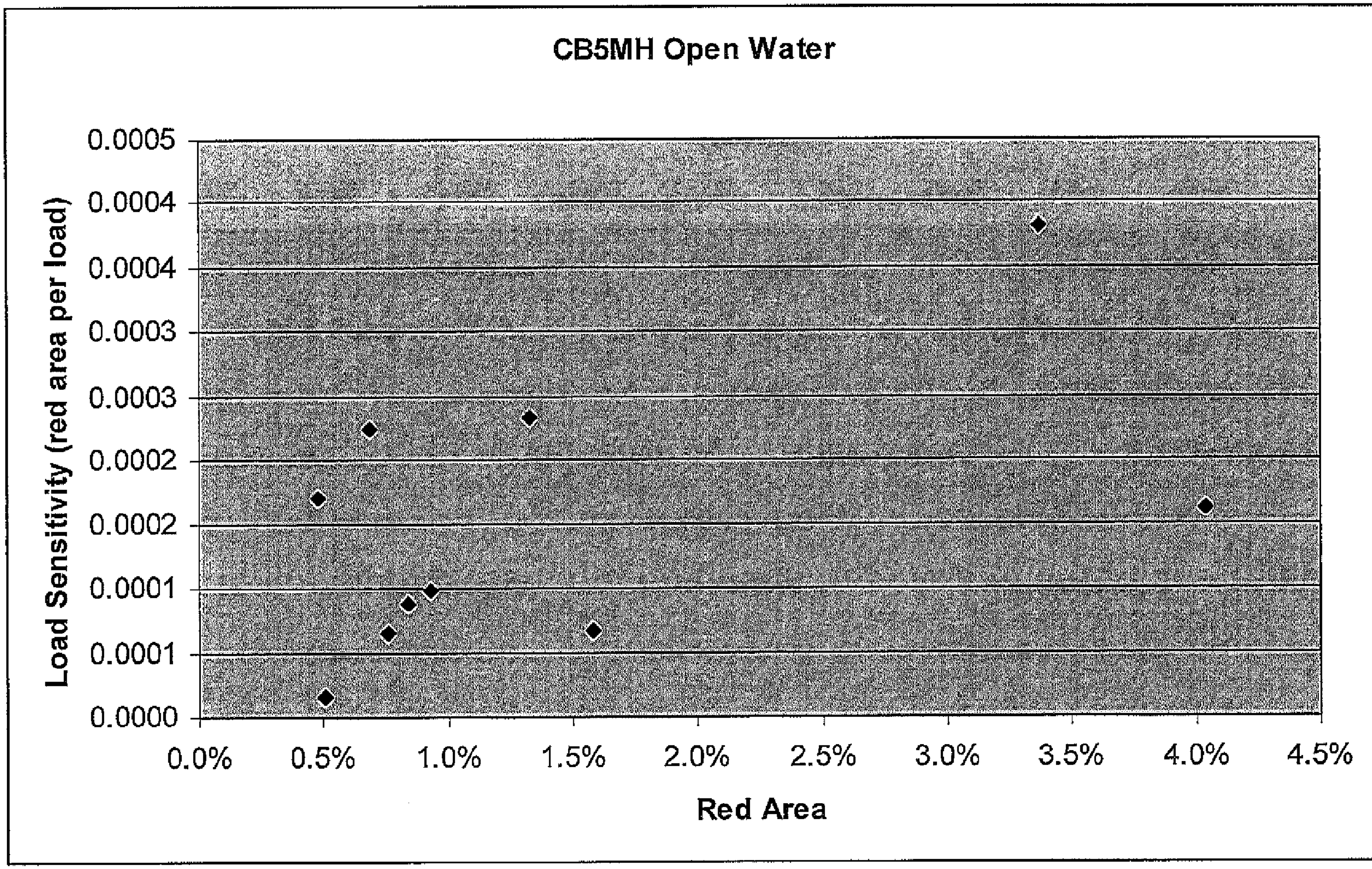
Sensitivity to Load Reductions

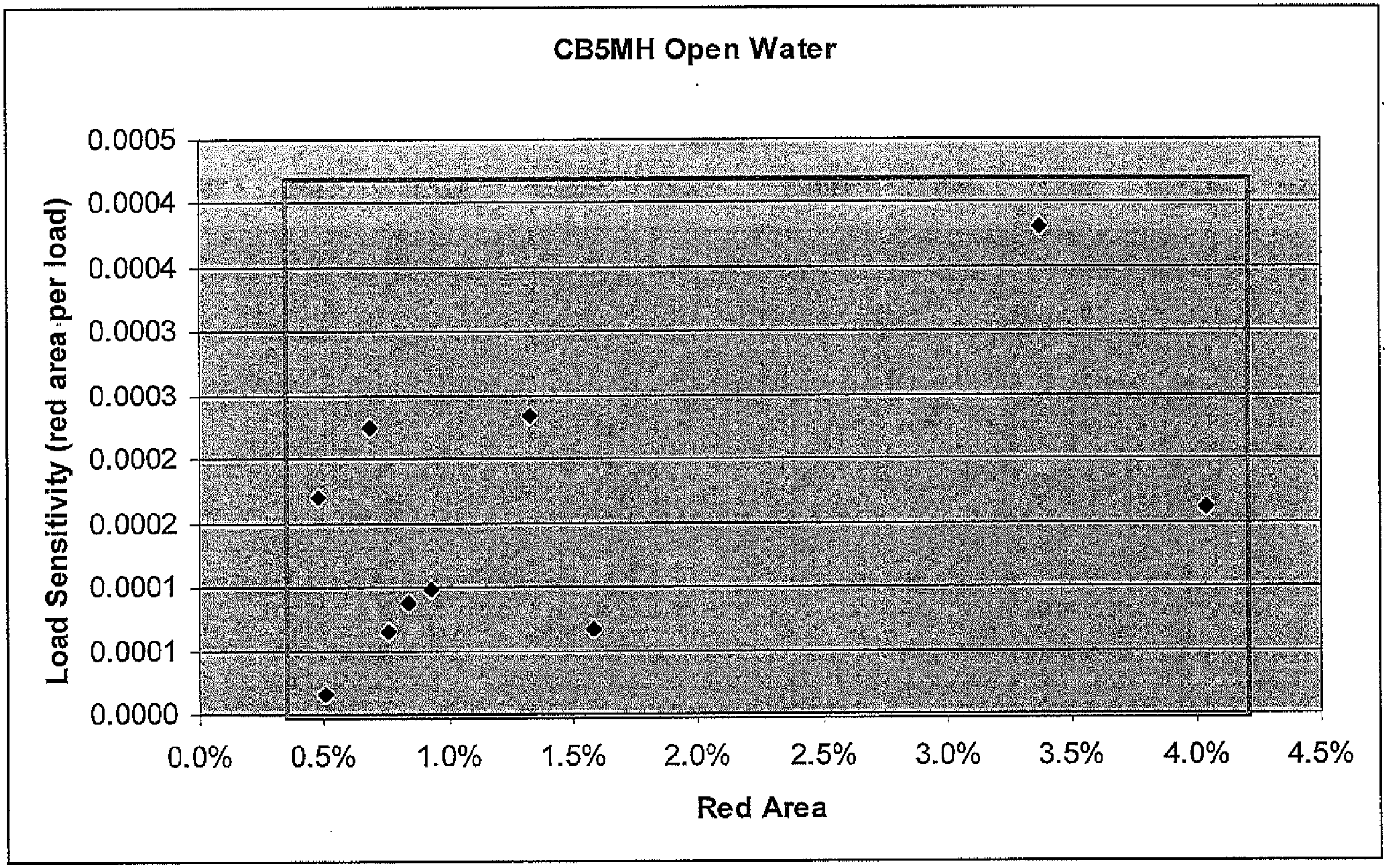
- Focused on segments driving the Bay TMDL
  - CB3MH, CB4MH, CB5MH for DW and DC and POTMH for DW
- Provided two open-water examples to show contrasts
  - CB2OH: drop in sensitivity at low non-attainment values
  - CB5MH: sensitivity to load reductions relatively constant throughout model simulated range

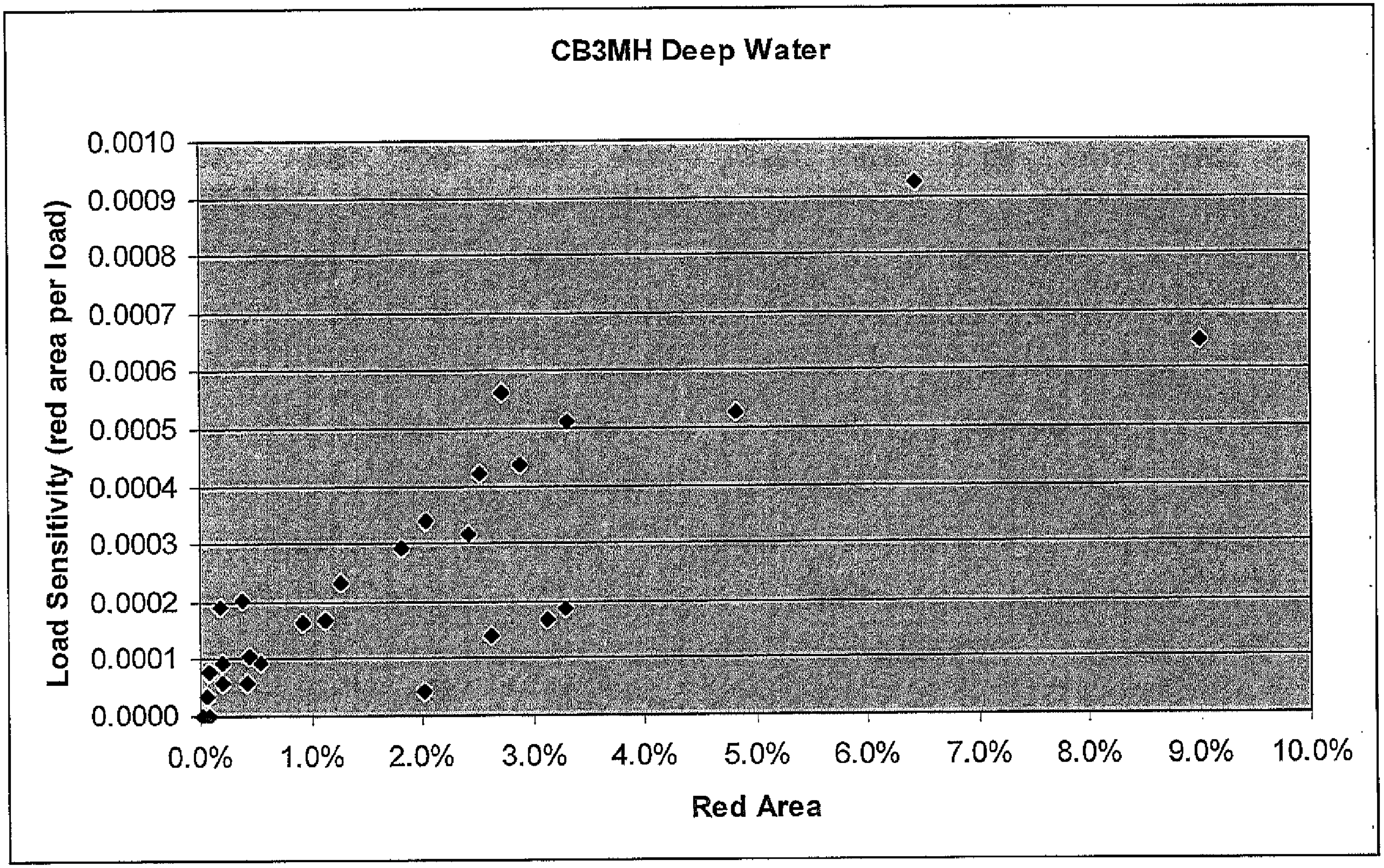
IC

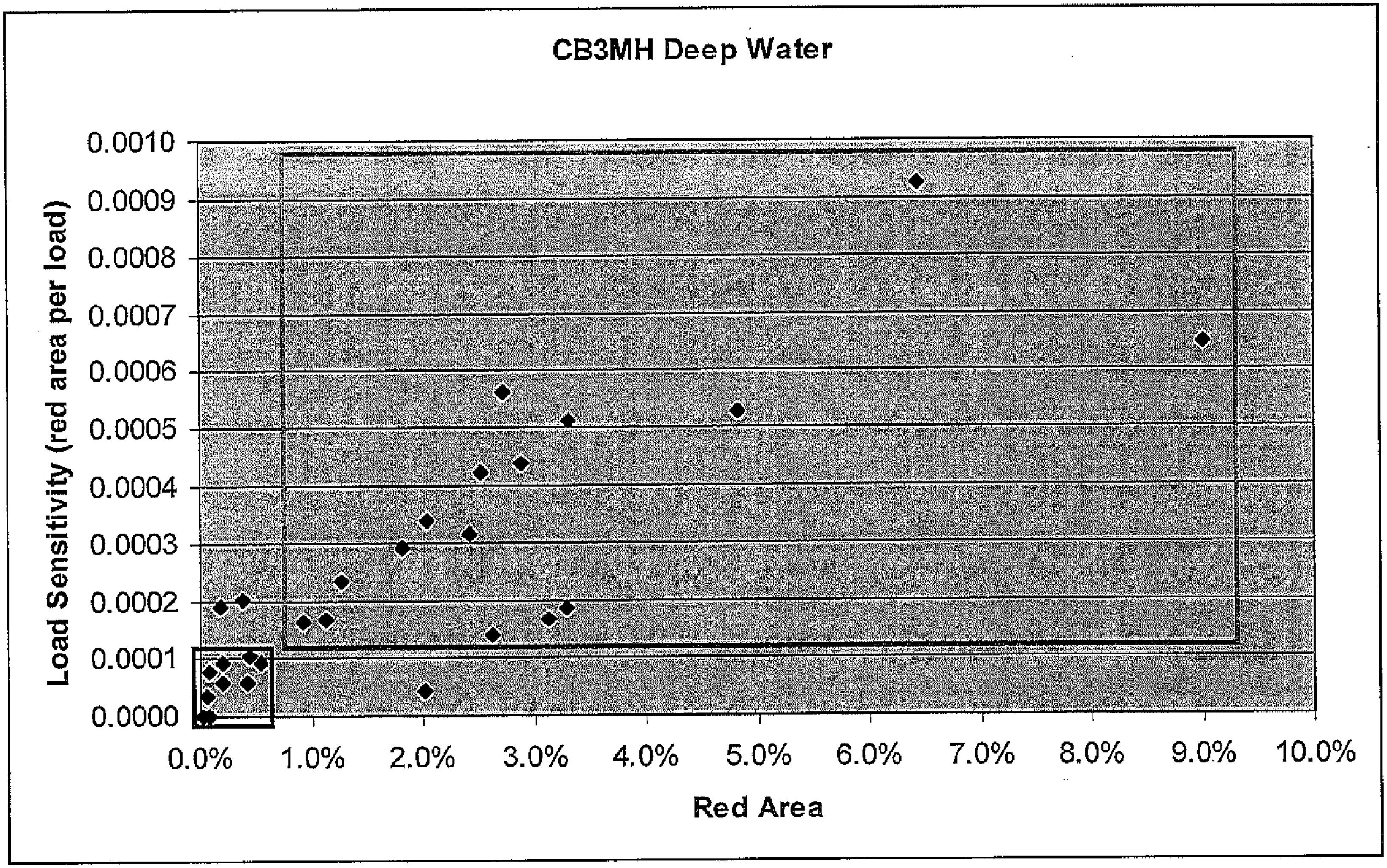

CB2OH Open Water
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
Load Sensitivity (red area per load)
0.0%
0.5%
1.0%
1.5%
2.0%
2.5%
3.0%
3.5%
4.0%
4.5%
5.0%
Red Area

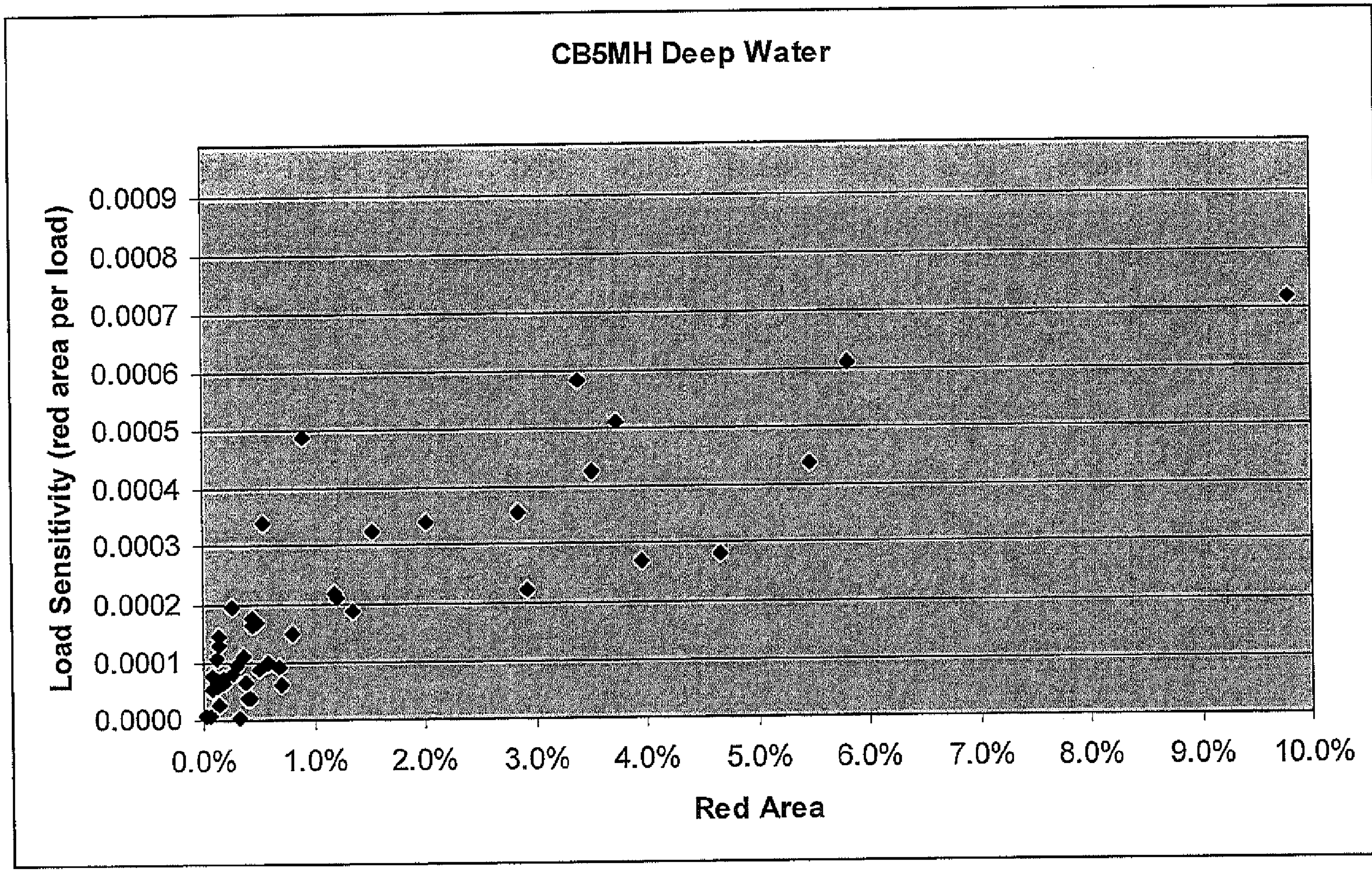

CB2OH Open Water
Load Sensitivity (red area per load)
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0.0%
0.5%
1.0%
1.5%
2.0%
2.5%
3.0%
3.5%
4.0%
4.5%
5.0%
Red Area

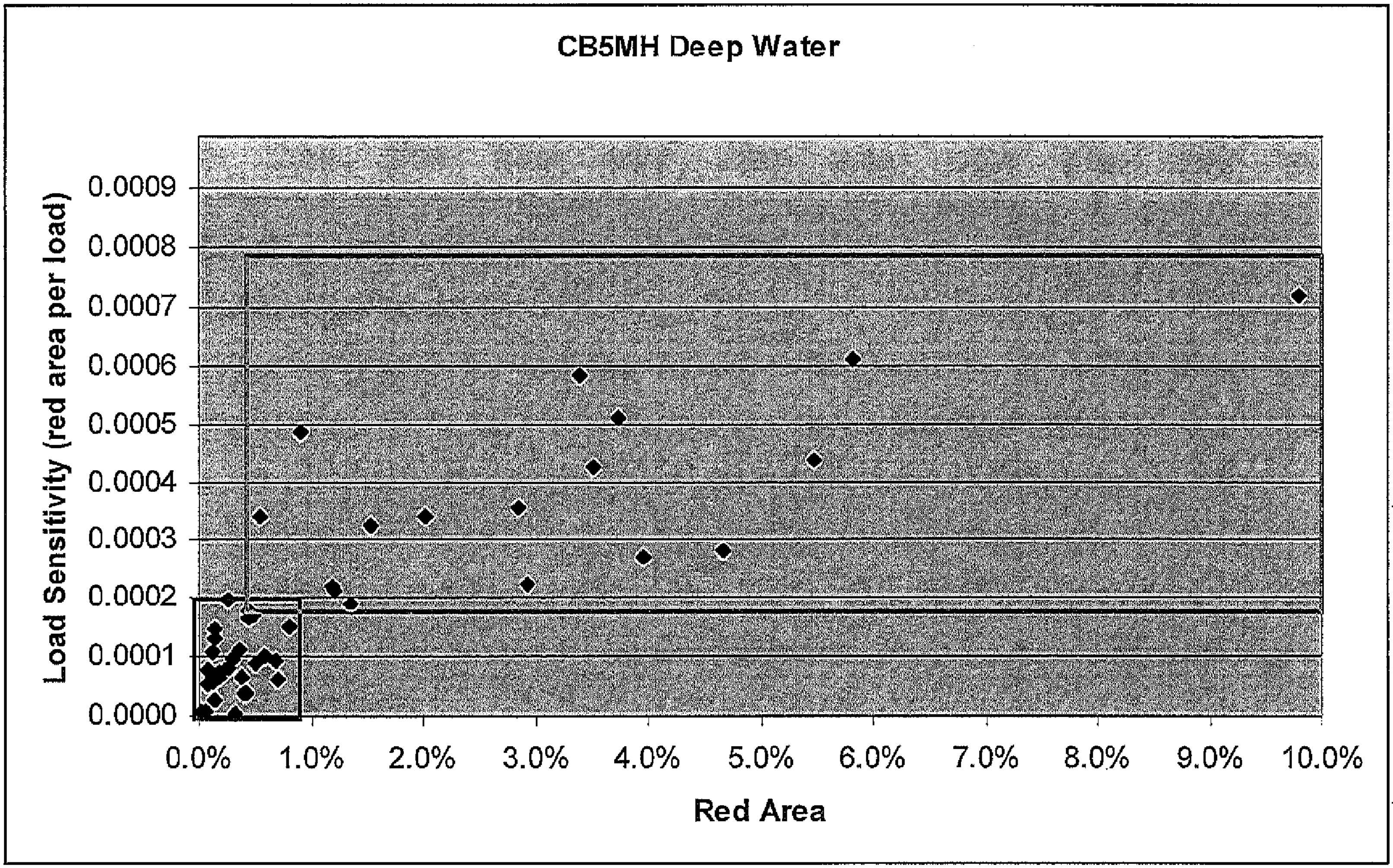

CB5MH Open Water
Load Sensitivity (red area per load)
0.0005
0.0004
0.0004
0.0003
0.0003
0.0002
0.0002
0.0001
0.0001
0.0000
0.0%
0.5%
1.0%
1.5%
2.0%
2.5%
3.0%
3.5%
4.0%
4.5%
Red Area

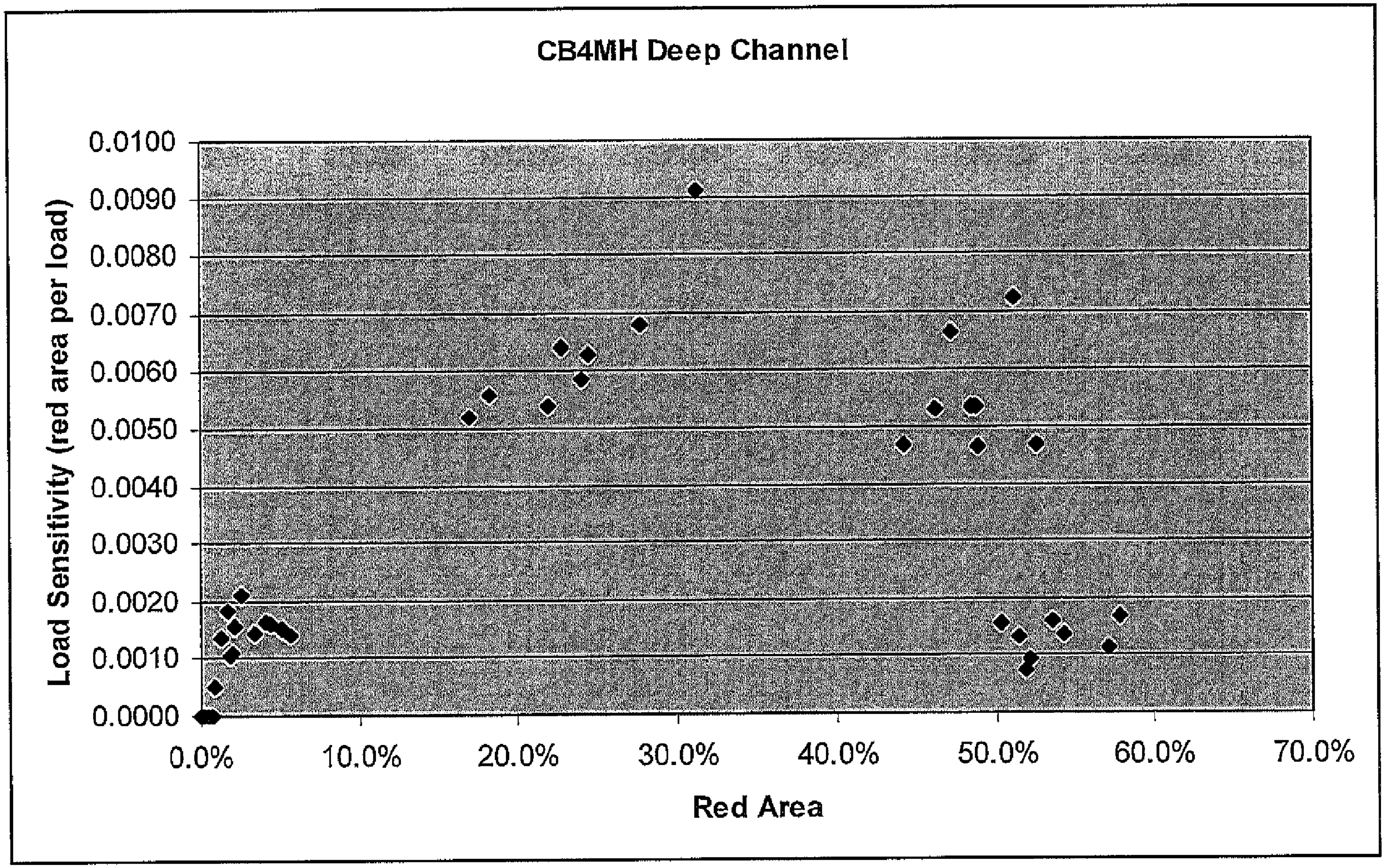

CB5MH Open Water
Load Sensitivity (red area per load)
0.0005
0.0004
0.0004
0.0003
0.0003
0.0002
0.0002
0.0001
0.0001
0.0000
0.0%
0.5%
1.0%
1.5%
2.0%
2.5%
3.0%
3.5%
4.0%
4.5%
Red Area

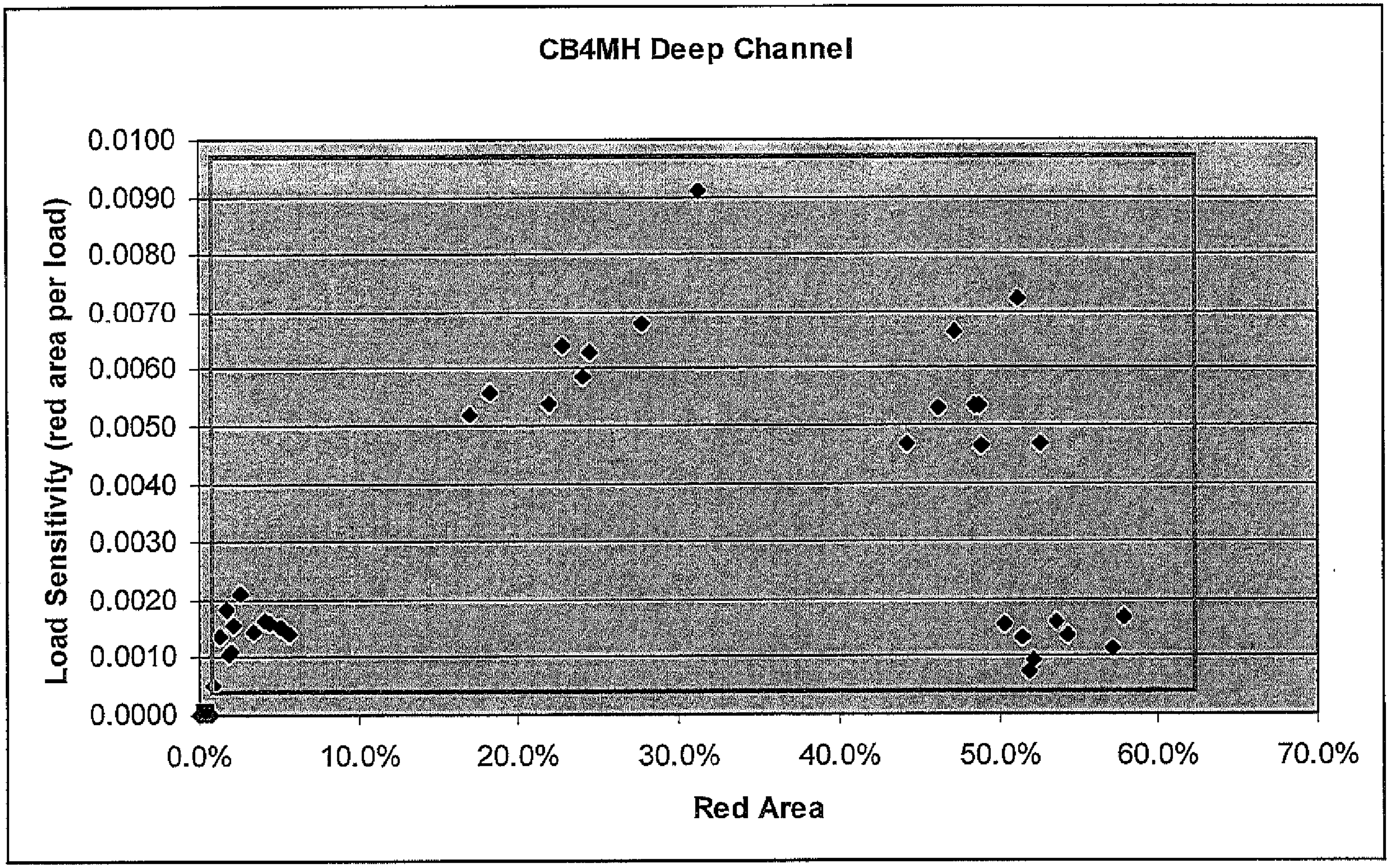

CB3MH Deep Water
Load Sensitivity (red area per load)
0.0010
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0.0%
1.0%
2.0%
3.0%
4.0%
5.0%
6.0%
7.0%
8.0%
9.0%
10.0%
Red Area

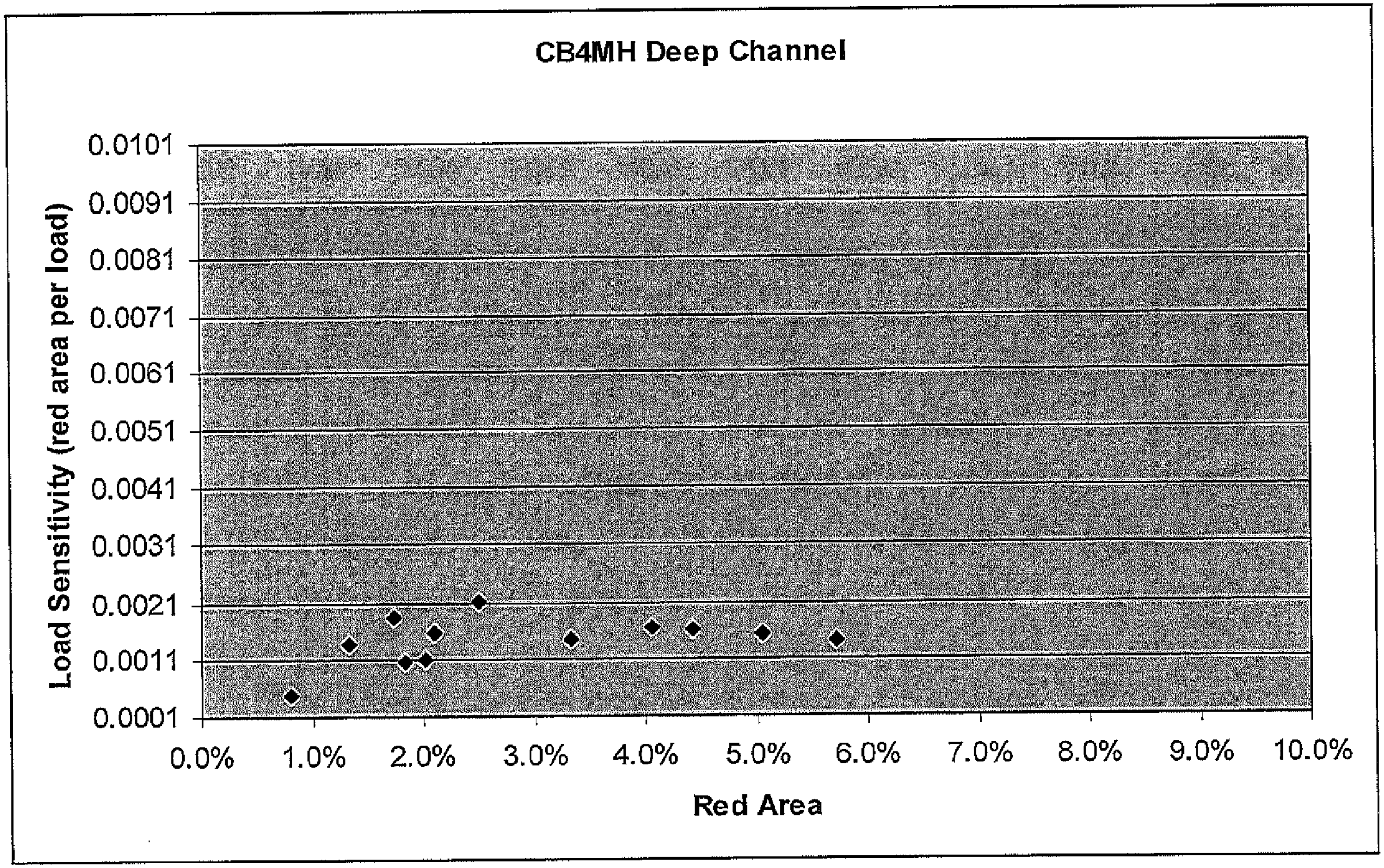

CB3MH Deep Water
Load Sensitivity (red area per load)
0.0010
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0.0%
1.0%
2.0%
3.0%
4.0%
5.0%
6.0%
7.0%
8.0%
9.0%
10.0%
Red Area

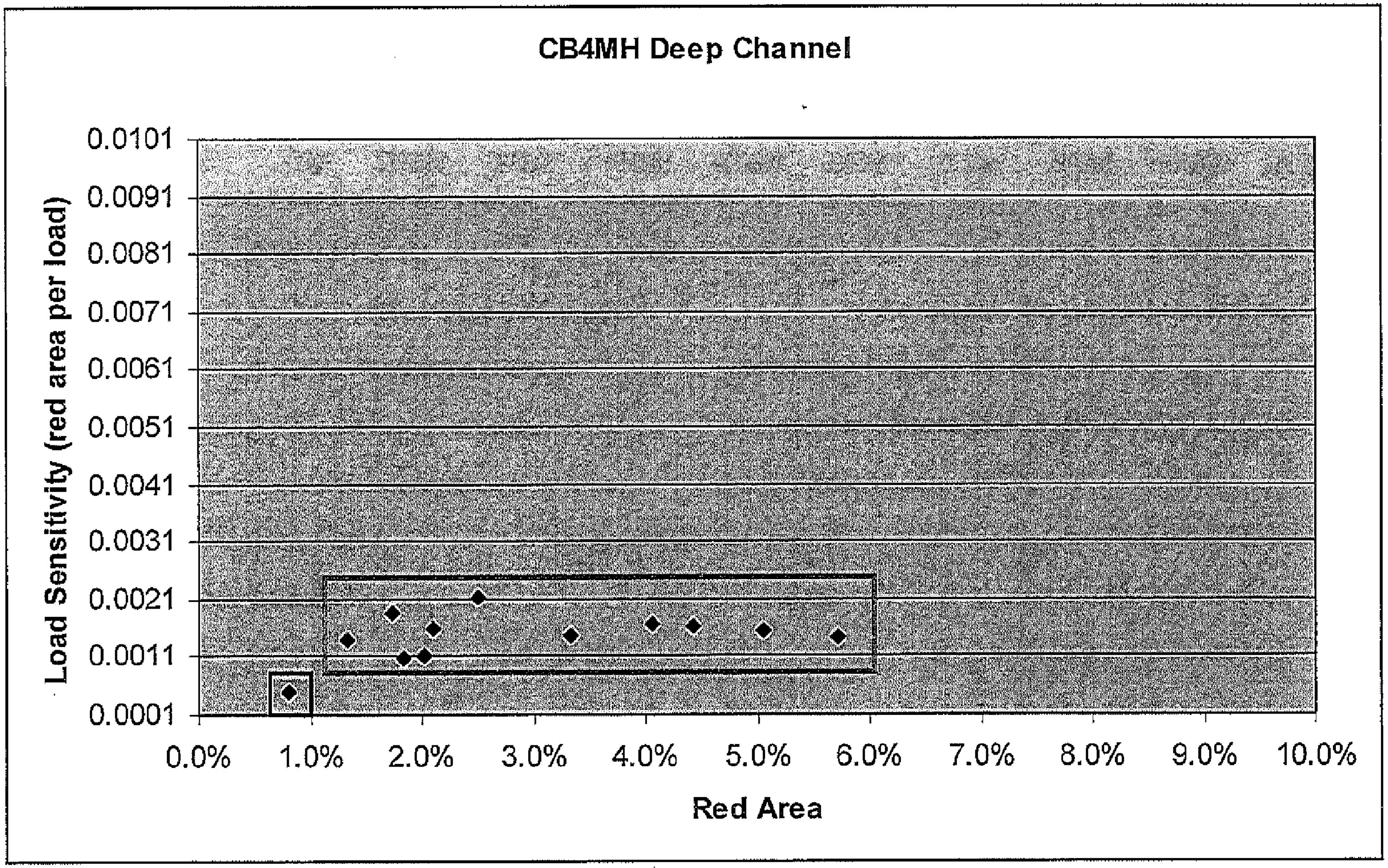

CB5MH Deep Water
Load Sensitivity (red area per load)
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0.0%
1.0%
2.0%
3.0%
4.0%
5.0%
6.0%
7.0%
8.0%
9.0%
10.0%
Red Area


CB5MH Deep Water
Load Sensitivity (red area per load)
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0.0%
1.0%
2.0%
3.0%
4.0%
5.0%
6.0%
7.0%
8.0%
9.0%
10.0%
Red Area


CB4MH Deep Channel
Load Sensitivity (red area per load)
0.0100
0.0090
0.0080
0.0070
0.0060
0.0050
0.0040
0.0030
0.0020
0.0010
0.0000
0.0%
10.0%
20.0%
30.0%
40.0%
50.0%
60.0%
70.0%
Red Area


CB4MH Deep Channel
Load Sensitivity (red area per load)
0.0100
0.0090
0.0080
0.0070
0.0060
0.0050
0.0040
0.0030
0.0020
0.0010
0.0000
0.0%
10.0%
20.0%
30.0%
40.0%
50.0%
60.0%
70.0%
Red Area


CB4MH Deep Channel
Load Sensitivity (red area per load)
0.0101
0.0091
0.0081
0.0071
0.0061
0.0051
0.0041
0.0031
0.0021
0.0011
0.0001
0.0%
1.0%
2.0%
3.0%
4.0%
5.0%
6.0%
7.0%
8.0%
9.0%
10.0%
Red Area


CB4MH Deep Channel
Load Sensitivity (red area per load)
0.0101
0.0091
0.0081
0.0071
0.0061
0.0051
0.0041
0.0031
0.0021
0.0011
0.0001
0.0%
1.0%
2.0%
3.0%
4.0%
5.0%
6.0%
7.0%
8.0%
9.0%
10.0%
Red Area

Chesapeake Bay Segment	Designated Use	Red Area with Low Sensitivity to Load Reductions (%)
CB3MH	Deep-water	0.2
CB4MH	Deep-water	0
CB5MH	Deep-water	1
POTMH	Deep-water	1
CB3MH	Deep-channel	1-1.5
CB4MH	Deep-channel	1
CB5MH	Deep-channel	N/A

Attainment at 1% Non-Attainment

- 21 designated use-segments with non-attainment values ranging from 0.0% to 1.5% across wide N load reductions

- Significant drop-off in sensitivity to load reductions documented at low non-attainment percentages
  - 1% being the consistent level at which sensitivity decreases across most of the designated use-segments driving the Bay TMDL

Attainment at 1% Non-Attainment

- The analysis system has been shown to be significantly less sensitive to the effects of further model simulated load reductions at and below the 1% non-attainment level

- These findings support documentation of attainment of designated use-segments with model simulated dissolved oxygen criteria non-attainment at 1%
  - No evidence documented for either analysis supporting a higher percentage (e.g., 2-3%)